EDITION 3

Business Statistics

ROBERT A. DONNELLY, JR.

With Special Contributions by

SERINA AL HADDAD, *Rollins College*

STEFAN RUEDIGER, *Arizona State University*

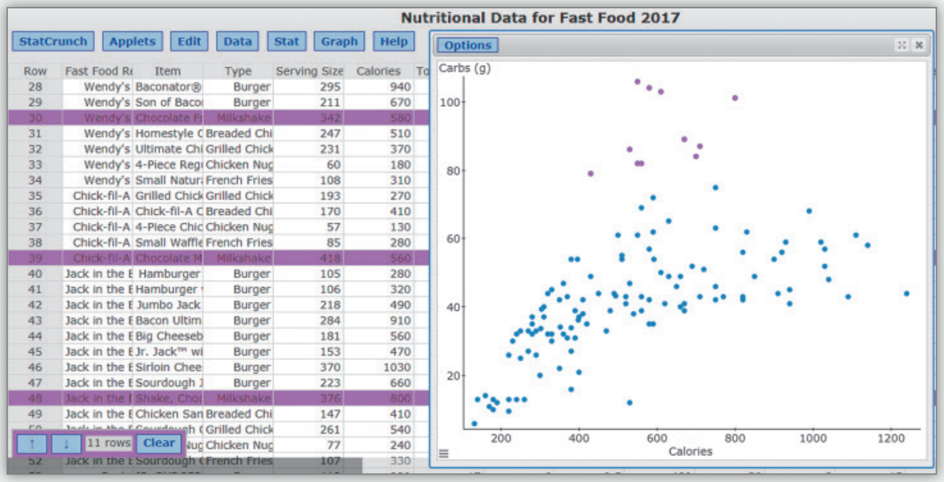
Get the Most Out of
MyLab Statistics

MyLab™ Statistics is the teaching and learning platform that empowers instructors to reach *every* student. By combining trusted author content with digital tools and a flexible platform, MyLab Statistics personalizes the learning experience and improves results for each student.

It makes learning and using statistical software seamless by allowing users to download and copy data sets directly into their programs. Students can access tutorial videos, as well as Study Cards for StatCrunch, Excel, Minitab, JMP, R, and SPSS. Many of these resources are also available within MyLab exercises as learning aids.

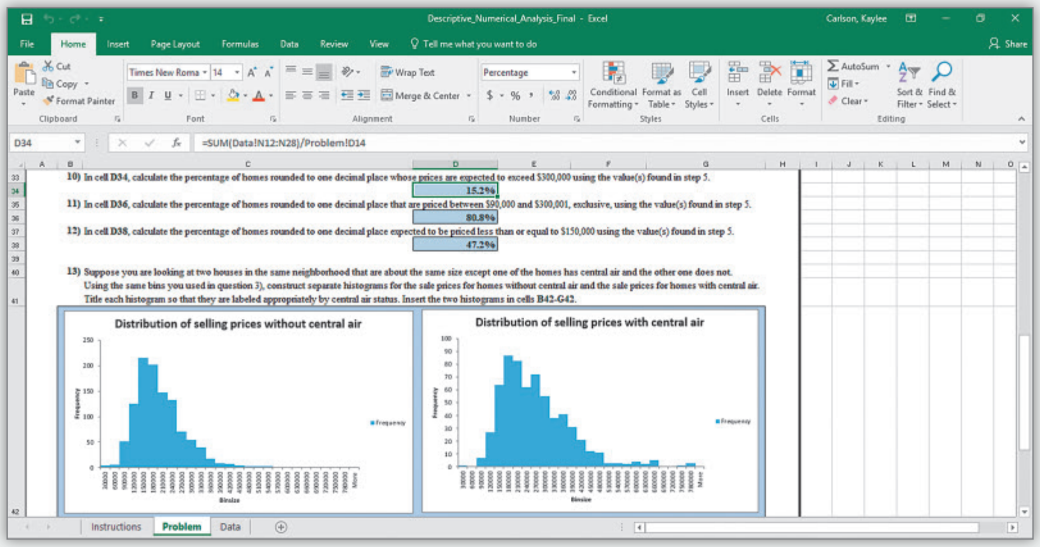
Collect, crunch, and communicate with StatCrunch

With StatCrunch®, Pearson's powerful web-based statistical software, access tens of thousands of data sets including those from the textbook, perform complex analyses, and generate compelling reports. Use StatCrunch directly in MyLab Statistics or as a standalone product.



Enjoy hands-off grading with Excel Projects

Using proven, field-tested technology, auto-graded Excel Projects let instructors seamlessly integrate Microsoft Excel content into the course without manually grading spreadsheets. Students can practice important statistical skills in Excel, use native Excel functions and Excel's Data analysis Add-in, helping them master key concepts and gain proficiency with the program.



Visit pearson.com/mylab/statistics and click Training & Support to make sure you're getting the most out of MyLab Statistics.

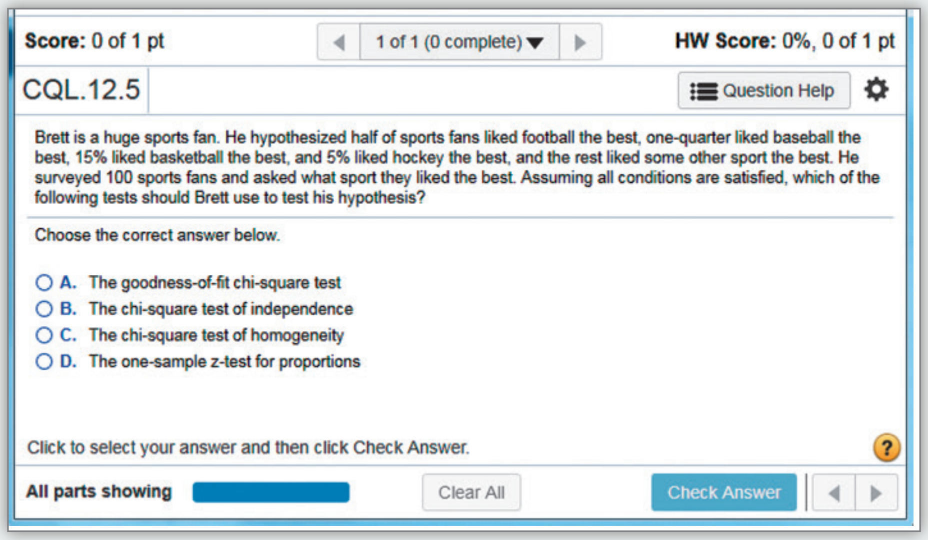
Pearson MyLab

Enrich assignments with question libraries

MyLab Statistics includes a number of question libraries providing additional opportunities for students to practice statistical thinking.

- StatCrunch Projects provide students with opportunities to analyze and interpret data. Each project consists of a series of assignable questions about a large data set.

- The Conceptual Question Library offers 1,000 conceptual-based questions to help students internalize concepts, make interpretations, and think critically about statistics.



- The Getting Ready for Statistics Library contains more than 450 exercises on prerequisite topics. Assign these questions to students who may need a little extra practice on their prerequisite skills to be successful in your course.

- The StatTalk Video Library is based on a series of 24 videos, hosted by fun-loving statistician Andrew Vickers, that demonstrate important statistical concepts through interesting stories and real-life events.

- Business Insight Videos show how managers at top companies use statistics in their every-day work.

Incorporate additional author-created resources in class

Authors infuse their own voice, approach, and experiences teaching statistics into additional text-specific resources, such as interactive applets, technology manuals, workbooks, and more. Check out the Preface to learn more about what's available for this specific title.

EDITION 3

Business Statistics

ROBERT A. DONNELLY, JR.

With Special Contributions by

SERINA AL HADDAD, *Rollins College*

STEFAN RUEDIGER, *Arizona State University*

Pearson

Director, Portfolio Management: Deirdre Lynch
Specialist Portfolio Management Analyst: Suzanna Bainbridge
Editorial Assistant: Morgan Danna
Managing Producer: Scott Disanno
Content Producer: Peggy McMahon
Senior Publishing Services Analyst: Joe Vetere
Manager Quality Control: Mary Durnwald
Manager Content Development: Robert Carroll
Producer: Jean Choe
Product Marketing Manager: Lindsey Erlick
Manager, Rights/Permissions: Gina Cheselka
Full Service Vendor: Pearson CSC
Full Service Project Management: Pearson CSC, Sharon Cahill
Composition: Pearson CSC
Text and Cover Design: Studio Montage
Cover Images: Stock market analysis data on computer screen: Nokwan007/Shutterstock; Row of cars: Jaggat Rashidi/ Shutterstock; Students Collaborating On Project On Campus: Monkey Business Images/Shutterstock; 3d printer, printing house Illustration: Belekekin/Shutterstock; Two cups of Coffee: Oleg Samoylov/Shutterstock.
Printer/Binder: LSC Communications, Inc.
Cover Printer: LSC Communications, Inc.

Credits and acknowledgments borrowed from other sources and reproduced, with permission, in this textbook appear on the appropriate page within text or on page 934 in the back of this book, which is hereby made part of this copyright page.

Library of Congress Cataloging-in-Publication Data
Names: Donnelly, Robert A., Jr., author.
Title: Business statistics / Robert A. Donnelly, Jr. ; with special contributions by Serina AL Haddad, Rollins College, Stefan Ruediger, Arizona State University.
Description: 3rd edition. | Boston : Pearson, [2020] | Includes bibliographical references and index.
Identifiers: LCCN 2018059450| ISBN 9780134688121 (instructor's edition) | ISBN 0134688120 (instructor's edition) | ISBN 9780134685267 (student edition) | ISBN 0134685261 (student edition)
Subjects: LCSH: Commercial statistics. | Industrial management--Statistical methods.
Classification: LCC HF1017 .D646 2020 | DDC 519.5--dc23 LC record available at https://lccn.loc.gov/2018059450

1 19

Pearson

Student Edition
ISBN-10: 0-13-468526-1
ISBN-13: 978-0-13-468526-7

To my wife, Debbie, who supported and encouraged me every step of the way. I could not have done this without you, Babe.

This page intentionally left blank

ABOUT THE AUTHOR

Robert A. Donnelly, JR.

The late ***Robert (Bob) A. Donnelly, Jr.,*** was a professor at Goldey-Beacom College in Wilmington, Delaware, with more than 25 years of teaching experience. He taught classes in statistics, operations management, spreadsheet modeling, and project management at both the undergraduate and graduate levels. Bob earned an undergraduate degree in chemical engineering from the University of Delaware, after which he worked for several years as an engineer with the Diamond Shamrock Corporation in a chlorine plant. Despite success in this field, Bob felt drawn to pursue a career in education. It was his desire to teach that took him back to school to earn his MBA and Ph.D. in Operations Research, also from the University of Delaware. Bob also taught in the MBA program at the International School of Management in Paris, France. He thoroughly enjoyed discussing research methods and business statistics with both his French and American students.

Bob's working experience prior to his teaching career provided him with many opportunities to incorporate real-life examples into classroom learning. His students appreciated his knowledge of the business world as well as his mastery of the course subject matter. Many former students have sought Bob's assistance in work-related issues that dealt with his expertise. Typical student comments focused on his genuine concern for their welfare and his desire to help them succeed in reaching their goals.

ABOUT THE SPECIAL CONTRIBUTORS

SERINA AL HADDAD

Serina Al Haddad, Ph.D., is a visiting assistant professor in the business department at Rollins College in Winter Park, Florida, where she teaches classes in business statistics and operations management. She taught classes in business statistics and analytics at Stetson University for three years and received the Outstanding Teacher of the Year Award in 2017. Serina also taught undergraduate and graduate classes in statistics and decision analysis at the University of Central Florida (UCF). She completed her doctorate at UCF in industrial engineering and management systems. She has more than fourteen years of experience working in both academia and industry. She held multiple positions in statistical quality control, customer relationship management, institutional development, and business process reengineering. Her research interests include statistics, business analytics, innovation, and organizational development.

Serina is a freelance consultant and trainer. She is one of the main authors of Excel projects in Pearson MyLab Statistics. She has worked with many organizations including Junior Achievement, State University of New York and the United States Agency for International Development (USAID). She is a member of the Institute of Industrial and Systems Engineering (IISE). She was a founding member and served as president of the IISE student chapter at her undergraduate university. Serina's working experience brings a lot of real-life examples to the classroom. Her students appreciate her passionate teaching, which keeps them highly engaged throughout the learning process.

STEFAN RUEDIGER

Stefan Ruediger, Ph.D., is a clinical associate professor in the economics department in the W. P. Carey School of Business at Arizona State University (ASU), where he teaches classes in business statistics, managerial economics, and principles of macroeconomics. He taught classes in economics at the University of Wisconsin–Stevens Point from 2008 to 2013 and received the School of Business and Economics Teaching Excellence Award in 2012. Stefan also taught undergraduate classes in economics at the University of Wisconsin–Parkside and at the University of Wisconsin–Milwaukee (UWM). He completed his doctorate at UWM. He has more than fifteen years of experience working in academia.

His research is focused on economic and statistics education, economic development, and financial economics. He has several publications in these areas and has won awards for his research. Stefan has worked in collaboration with colleagues at ASU on developing new adaptive courseware tools, in-class instructional materials, instructional videos, and teaching modes to improve student success, increase student retention rates, and decrease withdrawal rates. These projects have been supported by grants from the Gates Foundation and the Association of Public and Land-grant Universities (APLU).

In addition, Stefan was a track and field coach at UWM for five years and a member of the German Bobsledding national team for four years. His athletic background always provides a good basis for fun conversations with students and is a source of great economics and statistics examples.

BRIEF CONTENTS

This page intentionally left blank

TABLE OF CONTENTS

PREFACE

Business Statistics, ***third edition***, is a one- or two-semester textbook written in a conversational tone designed to reduce the level of anxiety that many students experience when taking a statistics course.

Many of today's business students are intimidated by their statistics textbook. These students often view their textbook as an obstacle to overcome rather than a tool to help them succeed. To address this issue, The late Robert Donnelly wrote *Business Statistics* in a straightforward, conversational tone that helps to reduce the anxiety many students experience with this course. His experience as both a writer and a teacher taught him that students learn more effectively when they feel a personal connection with their instructor. Many traditional textbooks tend to "talk down" to students in a manner that many find difficult to understand. Donnelly preferred a textbook that "talks to" the students as he did in the classroom, providing them a sense that he is on their side, encouraging them every step of the way.

We, Bob Donnelly and the two new contributors, Serina Al Haddad and Stefan Reudiger, strongly believe that students learn most effectively when they solve statistics problems as they learn new concepts rather than later (often right before the next exam). To facilitate this philosophy, we provide the student with a parallel problem labeled "Your Turn," which allows them to work alongside the example that is being demonstrated in the chapter. We attempt to motivate them to do these exercises with a little levity but it's not beneath us to downright beg them to give it a try. We show the entire solution at the end of the chapter, so they can quickly check if their answer is correct. We call this the "learn it, do it, check it" cycle, where students learn by reading an example, doing a similar problem on their own, and finally checking their answer to confirm they understand. In effect, the textbook also plays the role of a workbook for the students, keeping them actively engaged in the learning process. Too often, students skim through an example that is completely solved for them in the text and convince themselves they understand the concept—that is, until they are trying to solve a similar problem in an exam for the first time. Our approach encourages students to work through examples and confirm they grasp the concept before moving on to the next topic.

We have inserted many author's comments in the margins that provide useful insights along the way. This feature is analogous to the side comments you would make to your students during a lecture to help them better understand the material. We have found this to be an effective technique to help keep students focused on material that they may find confusing or overwhelming.

To help students be successful in your course, *Business Statistics*, third edition, has the following attributes:

- **Is written in a straightforward, conversational tone** to reduce the anxiety that many business students experience with the topic of statistics.
- **Uses a "learn it, do it, check it" cycle** by providing parallel Your Turn problems throughout each chapter, the textbook essentially serves as a workbook, allowing students to convince themselves they really understand a concept before moving on to the next topic.
- **Incorporates author's comments in the margins,** which are analogous to the side comments that an instructor would make during a lecture to help students better understand the material.
- **Utilizes technology (Microsoft Excel 2016)** in applying statistical analyses and refers students to available Excel projects, when applicable.

NEW TO THIS EDITION

We are very excited to offer several new features and many updated features in the third edition of *Business Statistics*. We have

- **Updated technology coverage** with Microsoft Excel 2016, including steps and figures. Excel as an analytical tool is not expensive and is often available on computers. Excel has become one of the most utilized analysis and business intelligence tools. This edition focuses on native Excel and Excel's Data Analysis ToolPak.

- **Updated data sets** in problems, examples, and Your Turns with current available data and replaced outdated or obsolete examples and data sets with more timely topics and data sets that relate to current business systems, technologies, and applications.
- **Focused more on Excel** and Excel's Data Analysis ToolPak, which covers most functionalities needed in this course.

e PROJECT: EXCEL

Describing_Data_A_Frequency_Distributions

- Added **Excel Projects** and callouts to these projects relating to a certain topic or section. The topics span the entire course, from describing and calculating data to hypothesis testing to linear regression. These Excel assignments allow students to apply statistical analysis using Excel native functions and Excel's Data Analysis Add-in.
- Added **Focus on Analytics,** a new feature, in Chapters 1 through 16. This feature discusses business analytics examples and applications, which is an area that uses statistical analysis in the business world to come to meaningful conclusions and make decisions. Business analytics has become one of the top management approaches for achieving greater competitiveness.

FOCUS ON ANALYTICS Stock Investment

We have learned about probability and that there is a great deal of uncertainty and risk when making decisions in the business world. Probability plays an important role in the stock market because decisions made to invest in stocks are built on the probability of the price of the stock going up or down.

Since it is not easy to know the probability of an increase in a certain stock price, many investors rely on the average of significant stocks on the market and act accordingly. For example, if an investor is interested in knowing if a particular stock's price will increase, the investor can use the Dow Jones Industrial Average, which reflects the average of 30 significant stocks traded on the New York Stock Exchange (NYSE) and NASDAQ.

If there is a relationship between the Dow Jones and your stock, then there is an intersection between the two events "stock price going up" and "Dow Jones going up," as shown in orange in Figure 4.11.

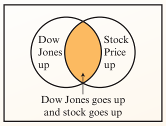


FIGURE 4.11 Conditional Probability in Stock Investment

Once the investor knows that the Dow Jones has gone up, he knows that no event outside the Dow Jones circle can occur. Since the attention is now restricted to "Dow Jones Up" circle, the investor needs to find the probability of the stock price going up within that restricted space, which is the area in orange. This is an application of conditional probability, which can be represented mathematically by:

$$P(\textit{Investor Stock goes up} \mid \textit{Dow Jones went up}) = \frac{P(\textit{Investor Stock goes up} \cap \textit{Dow Jones went up})}{P(\textit{Dow Jones went up})}$$

- In Chapter 1, introduced a topic on data sets and databases that explains definitions and shows examples and tables.
- In Chapter 2, added a topic that explains how to create histograms using Excel charts and a topic on Sparklines with definitions and steps. The polygon topic was replaced by the ogive.
- In Chapter 3, added a topic that explains short-cut formulas to find the sample variance and standard deviation in Excel in as well as a topic that demonstrates the use of Excel to draw box-and-whisker plots.
- In Chapter 4, added a topic that explains how to create contingency tables with probabilities in Excel by using a pivot table. More tables and explanations have been added for Bayes' theorem to simplify the logic and the process of deriving Bayes' theorem.
- In Chapter 5, added a topic that explains how to find the mean of a discrete probability distribution in Excel.
- In Chapter 9, added a topic that explains how to use Excel to perform a hypothesis test for the mean when σ is known.
- In Chapters 12 and 13, added a topic that explains how to use Excel to determine the p-value.

TEXTBOOK FEATURES

- **Current business examples that keep the students' interest**—Statistical procedures are applied to products and services that students can relate to such as the following:
  - Approximating the probability of an accident similar to the BP oil spill in the Gulf of Mexico occurring again in the near future
  - Liberty Mutual Insurance Company comparing the proportion of auto accident claims for clients with and without good student discounts
  - Comparing satisfaction scores for various smartphone brands
  - The shortage of Internet protocol (IP) addresses using the original IPv4 format
- **Your Turn problems after every major section**—These problems are strategically placed throughout the chapter and are designed to reinforce key concepts. Complete solutions to these problems can be found at the end of the chapter. Students learn more effectively when they actively solve problems rather than skim through examples that are completely solved for them.

> 4 **YOUR TURN**
>
> **1.** The following data represent the number of books that seven adults have read during the past 12 months. Determine the variance and standard deviation of these data.
>
> 10 10 4 8 13 6 11
>
> **2.** The ages of eight employees at a local PNC Bank branch are shown here. Determine the variance and standard deviation of these data.
>
> 32 40 27 29 34 38 46 24
>
> Answers can be found on **page 144.**

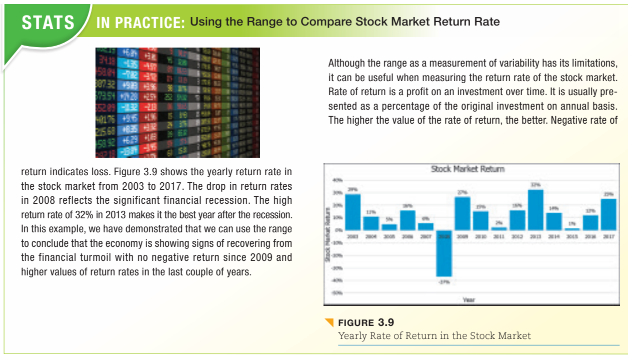
- **Step-by-step approach to complicated statistical procedures**—Many students tend to get "lost in the forest" when facing complicated procedures such as hypothesis testing, analysis of variance (ANOVA), and regression. We break these procedures down into bite-size, repeatable steps that can be applied to solve a variety of problems. As a result, the student has a consistent road map to follow when deciding how to proceed with a specific technique.
- **Highlighted text**—Throughout the textbook, we highlighted text to draw the student's attention to a key point in the chapter. This will reinforce important concepts that could otherwise be overlooked by the student.

> Even though the empirical probability of a tails in this example is 100%, we all know the classical probability is 50%. If we flipped the coin, say, 500 times, the empirical probability would be very close to 50%.

> When you rely on the mean to summarize many data values, you lose important information about the original data.

- **Author's comments**—In the margins, you will find comments that help clarify specific topics. These comments often point to an appropriate location in the chapter and are analogous to the comments you as an instructor might make in class to provide your students with some additional insight to the material.

- **Statistics in Practice**—Throughout the textbook, examples of how statistics is used in today's business environment are described in specially marked sections within the chapter. Examples include the following:
  - Government reports of unemployment figures using confidence intervals and the interpretation of these results
  - Comparison of customer feedback for different snack food products from Herr's Food Company
  - A statistical technique that health care insurance companies can rely on to investigate unusual billing practices from doctors' offices
  - Comparison of performances of Olympic athletes across different sporting events



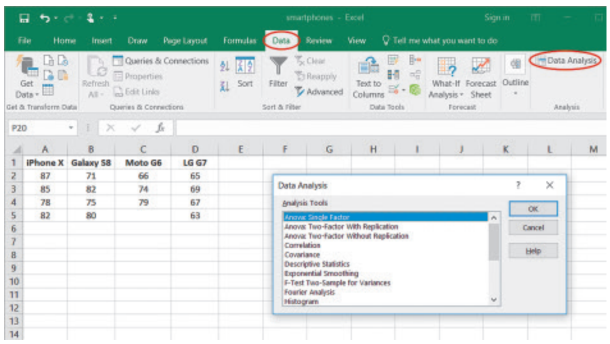
STATS IN PRACTICE: Using the Range to Compare Stock Market Return Rate

Although the range as a measurement of variability has its limitations, it can be useful when measuring the return rate of the stock market. Rate of return is a profit on an investment over time. It is usually presented as a percentage of the original investment on annual basis. The higher the value of the rate of return, the better. Negative rate of return indicates loss. Figure 3.9 shows the yearly return rate in the stock market from 2003 to 2017. The drop in return rates in 2008 reflects the significant financial recession. The high return rate of 32% in 2013 makes it the best year after the recession. In this example, we have demonstrated that we can use the range to conclude that the economy is showing signs of recovering from the financial turmoil with no negative return since 2009 and higher values of return rates in the last couple of years.

FIGURE 3.9
Yearly Rate of Return in the Stock Market

- **The integration of Microsoft Excel® 2016**—We use Excel to demonstrate the use of technology in business statistics, but not at the expense of understanding the underlying concepts. We have spoken to students who have said that they know how to perform a procedure such as ANOVA on the computer but do not feel comfortable interpreting the results. By letting Excel do all the work, they miss the opportunity to understand the underlying concepts of the technique. The philosophy employed in this textbook avoids this unfortunate situation.

FIGURE 11.8A
Conducting a One-Way ANOVA Test Using Excel (Steps 1–3)



e PROJECT: EXCEL
Describing_Data_Graphically_and_Numerically

- **Excel Project callouts**—After covering a topic or section, you will find callouts to Excel projects that are available in MyLab Statistics. These Excel projects allow you to apply the statistical analysis learned in an Excel file that you download and work on by following specific instruction, then upload for auto grading.

Get the Most Out of MyLab Statistics

Pearson MyLab

Statistics courses are continuously evolving to help today's students succeed. It's more challenging than ever to support students with a wide range of backgrounds, learner styles, and math anxieties. The flexibility to build a course that fits instructors' individual course formats—with a variety of content options and multimedia resources all in one place—has made MyLab Statistics the market-leading solution for teaching and learning mathematics since its inception.

78% of students say MyLab Statistics helped them learn their course content.*

Teach your course with a consistent author voice

With market-leading author content options, your course can fit your style. Pearson offers the widest variety of content options, addressing a range of approaches and learning styles, authored by thought leaders across the business and statistics curriculum. MyLab™ Statistics is tightly integrated with each author's style, offering a range of author-created multimedia resources, so your students have a consistent experience.

Thanks to feedback from instructors and students from more than 10,000 institutions, MyLab Statistics continues to transform—delivering new content, innovative learning resources, and platform updates to support students and instructors today and in the future.

**Source: 2018 Student Survey, n 31,721*

Pearson MyLab

Resources for Success

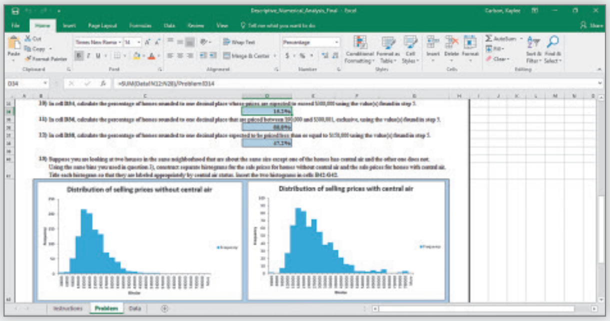
MyLab Statistics Online Course for *Business Statistics,* Third Edition, by Robert A. Donnelly

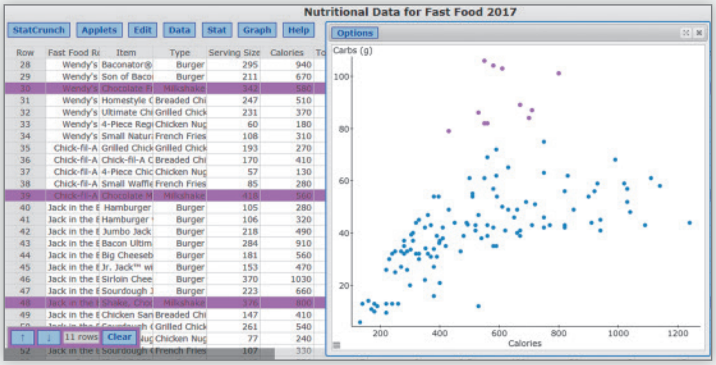
(access code required)

MyLab™ Statistics is available to accompany Pearson's market leading text offerings. To give students a consistent tone, voice, and teaching method, each text's flavor and approach are tightly integrated throughout the accompanying MyLab Statistics course, making learning the material as seamless as possible.

Enjoy hands-off grading with Excel Projects

Using proven, field-tested technology, auto-graded Excel Projects let instructors seamlessly integrate Microsoft Excel content into the course without manually grading spreadsheets. Students can practice important statistical skills in Excel, helping them master key concepts and gain proficiency with the program.



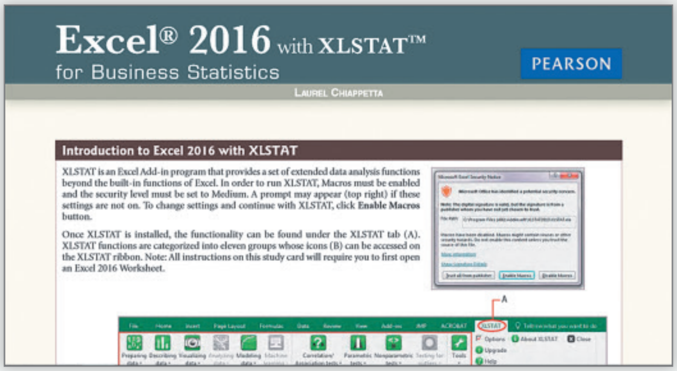


StatCrunch

Via StatCrunch, Pearson's powerful web-based statistical software, instructors and students can access tens of thousands of data sets, including those from the textbook, perform complex analyses, and generate compelling reports. StatCrunch is integrated directly into MyLab Statistics or available as a standalone product. To learn more, go to www.statcrunch.com on any laptop, tablet, or smartphone.

Technology Tutorials and Study Cards

MyLab makes learning and using a variety of statistical software programs as seamless and intuitive as possible. Download data sets from the text and MyLab exercises directly into Excel. Students can also access instructional support tools including tutorial videos, study cards, and manuals for a variety of statistical software programs including StatCrunch, Excel, Minitab, JMP, R, and SPSS.



pearson.com/mylab/statistics

Pearson MyLab

Resources for Success

Instructor Supplements

Online resources can be downloaded from **www.pearsonhighered.com/mathstatsresources.**

Instructor's Edition contains answers to all exercises. (ISBN-13: 978-0-13-468812-1; ISBN-10: 0-13-468812-0

Online Test Bank (download only), includes true/false, multiple choice, fill-in, and problem-solving questions based on the definitions, concepts, and ideas developed in each chapter of the text. Available to qualified instructors through Pearson's online catalog at **www.pearson.com/us/higher-education** or within MyLab Statistics.

Instructor's Solutions Manual (download only), Dirk Tempelaar, contains detailed solutions to all of the exercises. The Instructor's Solutions Manual is available to qualified instructors through Pearson's online catalog at **www.pearson.com/us/higher-education or within MyLab Statistics.**

TestGen® Computerized Test Bank
(www.pearsoned.com/testgen) enables instructors to build, edit, print, and administer tests using a computerized bank of questions developed to cover all the objectives of the text. TestGen is algorithmically based, allowing instructors to create multiple but equivalent versions of the same question or test with the click of a button. Instructors can also modify test bank questions or add new questions. The software and test bank are available for download from Pearson's online catalog at **www.pearson.com/us/higher-education.** Test Forms (download only) are also available from the online catalog.

PowerPoint Lecture Slides
Free to qualified adopters, this classroom lecture presentation software is geared specifically to the sequence and philosophy of *Business Statistics*. Key graphics from the book are included to help bring the statistical concepts alive in the classroom. These files are available to qualified instructors through Pearson's online catalog at **www.pearson.com/us/higher-education** or within MyLab Statistics.

Student Resources

Business Statistics, for-sale student edition. (ISBN-13: 978-0-13-468526-7 ISBN-10: 0-13-468526-1

Student's Solutions Manual Dirk Tempelaar, provides detailed, worked-out solutions to odd-numbered exercises. (ISBN-13: 978-0-13-468612-7; ISBN-10: 0-13-468612-8

Study Cards for Business Statistics Software: This series of study cards, available for Excel 2016 with DAT: 0-13-457679-9; Excel 2016 with XLSTAT: 0-13-457683-7; StatCrunch: 0-13-397513-4, R: 0-13-522870-0; JMP: 0-321-64423-9; Minitab 17: 0-13-457681-0; SPSS: 0-32-164422-0; and R Studio: 0-13-522869-7 provides students with easy step-by-step guides to popular statistics software.

Excel Projects: These projects are available in MyLab Statistics and allow students to apply statistical concepts using Excel Native functions and Excel's Data Analysis add-in. They cover areas in descriptive and inferential statistics. Students download and work on an Excel file by following specific instructions, then upload it for auto grading. Students can review their submission for detailed feedback.

XLSTAT™ for Pearson Used by leading businesses and universities, XLSTAT is an Excel® add-in that offers a wide variety of functions to enhance the analytical capabilities of Microsoft Excel, making it the ideal tool for your everyday data analysis and statistics requirements. XLSTAT is compatible with all Excel versions. Available for bundling with the text as ISBN-13:978-0-321-75940-5; ISBN-10: 0-321-75940-0. Standalone: ISBN-13: 978-0-321-75932-0; ISBN-10: 0-321-75932-X

ACKNOWLEDGMENTS

Special Thanks

Thanks are due to the following people for their assistance and encouragement to the preparation of this edition: A special thanks goes to Deirdre Lynch and Suzanna Bainbridge for their support and guidance with this third edition; Dana Bettez for her development editing and keeping the manuscript development process on schedule; Thomas Hayward for promoting this edition with his marketing skills; Joe Vetere for converting all of the Excel screenshots to prepare them for reproduction in the textbook; Peggy McMahon for directing the content production process; Sharon Cahill for acting as project manager with SPi Global, the vendor resource for copyediting, art rendering, page composition, proofreading, and indexing; Morgan Danna for coordinating the reviewers; and Cathy Zucco-Teveloff and Gary Bliss for their accuracy checking to ensure an error-free textbook.

This textbook was shaped and contributed to by dozens of class testers, consultants, focus group participants, and reviewers who gave generously of their time, expertise, and creativity.

REVIEWERS FOR THE THIRD EDITION

Rosemary Birkel Wilson, *Washtenaw Community College*
Hannah Bolte, *Indiana University Bloomington*
Dinna M. Cichocki, *University at Buffalo*
André Maharaj, *Florida International University*
Lee McClain, *Western Washington University*
Rosemary Milewski Danaher, *Sacred Heart University*
Ignacio Molina, *Arizona State University*
Carol Putman, *Pennsylvania State University - Behrend College*
Ping Wang, *James Madison University*

Participants/Reviewers for the First and Second Editions

ALABAMA

Scott Bailey, *Troy University*

ARIZONA

Henry Ander, *Arizona State University*
Jason Bronowitz, *Arizona State University*
Linda Chattin, *Arizona State University*
Ashley Jacobson, *Chandler Gilbert Community College*
Nicolas Rouse, *Phoenix College*
Susan Sandblom, *Scottsdale Community College*
Yvonne M. Sandoval, *Pima Community College*

ARKANSAS

Tony Hunnicutt, *Ouachita Technical College*

CALIFORNIA

Asatar Bair, *City College of San Francisco*
Michael Brady, *California State University, Dominguez Hills*
Min Li, *California State University, Sacramento*
Khosrow Moshirvaziri, *California State University, Long Beach*
Christopher O'Byrne, *San Diego State University*
Ozgur Ozluk, *San Francisco State University*
Hamid Pourmohammadi, *California State University—Dominquez Hills*
Hindupur Ramakrishna, *University of Redlands*
Sunil Sapra, *California State University, Los Angeles*
Mark Tendall, *Stanford University*
Ruben Veliz, *Marymount California University*

COLORADO

Mark Dahlke, *Colorado State University*
Dale DeBoer, *University of Colorado, Colorado Springs*
Eric Huggins, *Fort Lewis College*
Rutilio Martinez, *University of Northern Colorado*
Alexandre Probst, *Colorado Christian University*
Richard Turley, *Colorado State University*

CONNECTICUT

Matt Rafferty, *Quinnipiac University*

DELAWARE

Roman Erenshteyn, *Goldey-Beacom College*
James Ford, *University of Delaware*
Angela Mitchell, *Wilmington College*

FLORIDA

Dipankar Basu, *Miami University*
James Mirabella, *Jacksonville University*
Gary Smith, *Florida State University*

GEORGIA

Kim Gilbert, *University of Georgia*
Latika Lagalo, *Emory University*
Katherine McClain, *University of Georgia*
Komanduri Murty, *Fort Valley State University*
Hilde Patron-Boenheim, *University of West Georgia*

ILLINOIS

Harold Beck, *Southern Illinois University Edwardsville*
Yanli Cui, *University of Illinois at Chicago*
Jian Li, *Northeastern Illinois University*

INDIANA

Jason Davidson, *Butler University*
Ellen Gundlach, *Purdue University*
Keith Starcher, *Indiana Wesleyan University*

IOWA

Rick Jerz, *St. Ambrose University*
Blake Whitten, *University of Iowa*

KENTUCKY

Bob Gillette, *University of Kentucky*
Kusum Singh, *University of Kentucky*
Meenu Singh, *Murray State University*

LOUSIANA

Brandi Guidry, *University of Lousiana Lafayette*
Begona Perez-Mira, *Northwestern State University of Louisiana*
Zhiwei Zhu, *University of Louisiana at Lafayette*

MARYLAND

John Beyers, *University of Maryland–University College*
Michael Kulansky, *University of Maryland—University College*
Elkanah Faux, *Bowie State University*
Steve South, *Montgomery College, Germantown*

MASSACHUSETTS

Richard McGowan, *Boston College*
Bharatendra Rai, *UMASS–Dartmouth*

MICHIGAN

Fred Gruhl, *Washtenaw Community College*

MINNESOTA

Kristian Braekkan, *Gustavus Adolphus College*

MISSISSIPPI

Daniel Cameron Montgomery, *Delta State University*

Stephen Trouard, *Mississippi College*

MISSOURI

Eli Snir, *Washington University in St. Louis*

NEBRASKA

Karl Petersen, *Bellevue University*

Jena Shafai, *Bellevue University*

NEW HAMPSHIRE

Irwin Bramson, *New Hampshire College*

NEW JERSEY

Emma Bojinova, *Canisius College*

Ellen Benowitz, *Mercer County Community College*

Raed Dandan, *William Paterson University*

Martin Gritsch, *William Paterson University*

David Letcher, *The College of New Jersey*

Lan Nygren, *Rider University*

Alex Olbrecht, *Ramapo College of New Jersey*

Penina Orenstein, *Seton Hall University*

Leonard Presby, *William Paterson University*

Harold Schneider, *Rider University*

NEW MEXICO

Sue Stockly, *Eastern New Mexico University*

NEW YORK

Kevin Caskey, *SUNY–New Paltz*

Ann Brandwein, *Baruch College*

Chun-Pin Hsu, *York College, City University of New York*

Jan Pitera, *Broome Community College*

Sean Simpson, *Westchester Community College*

NORTH CAROLINA

Alexander Deshkovski, *North Carolina Central University*

Casey DiRienzo, *Elon University*

Mahour Parast, *University of North Carolina–Penbroke*

OHIO

Pam Boger, *The Ohio State University*

Susan Emens, *Kent State University at Trumbull*

Eugene Jones, *The Ohio State University*

Joe Nowakowski, *Muskingum University*

Eddy Patuwo, *Kent State University*

Roxana Postolache, *Capital University*

Deborah Rumsey-Johnson, *The Ohio State University*

Anthony Sterns, *Kent State University, University of Maryland—University College*

Robert Stoll, *Cleveland State University*

Michael Welker, *Franciscan University of Steubenville*

Ronald Young, *Stark State College*

OKLAHOMA

Ramona Piearcy, *Connors State College*

Allen White, *Bacone College*

Diana Wolfe, *Oklahoma State University at Oklahoma City*

OREGON

Candace Sorenson, *Marylhurst University*

PENNSLYVANIA

Mohamed Albohali, *Indiana University of Pennsylvania*

Ozgun Caliskan Demirag, *Pennsylvania State-Behrend*

Deborah Gougeon, *University of Scranton*

Ian Langella, *Shippensburg University of Pennsylvania*

Edward Mathis, *Villanova University*

Jerrold H. May, *KGSB/University of Pittsburgh*

Robert O. Neidigh, *Shippensburg University*

Robert Patterson, *Pennsylvania State-Erie*

Carol Puttman, *Pennsylvania State-Erie*
Rick Tannery, *Slippery Rock University*
Ray Venkataraman, *Pennsylvania State–Behrend*
Charles Wilf, *Duquesne University*

SOUTH CAROLINA

Joan Donohue, *University of South Carolina*
Lisa Bosman, *Clemson University*
Chen-Huei Chou, *College of Charleston*
Kent Foster, *Winthrop University*
Renu Singh, *South Carolina State University*

TENNESSEE

Mohammad Ahmadi, *The University of Tennessee at Chattanooga*
Michael Racer, *University of Memphis*
Carolyn Rochelle, *East Tennessee State University*
Jeffrey Schultz, *Christian Brothers University*

TEXAS

Jacob Dell, *University of Texas at San Antonio*
Levon R. Hayrapetyan, *Houston Baptist University*
Levent Kaan, *University of Texas–Dallas*
Mark Leung, *University of Texas at San Antonio*
Carolyn Monroe, *Baylor University*
Ranga Ramasesh, *Texas Christian University*
Joseph C. Rhodes Jr., *Lone Star College*
Jesus Tanguma, *The University of Texas–Pan American*
Grace Vaughn, *El Paso Community College*

UTAH

Camille Fairbourn, *Utah State University Brigham City Regional Campus*
Don Gren, *Salt Lake Community College*

VIRGINIA

Kelly Alvey, *Old Dominion University*
Niels-Hugov Blunch, *Washington and Lee University*
Stephen Custer, *Virginia Commonwealth University*
Drew Koch, *James Madison University*
Quinton Nottingham, *Virginia Polytechnic Institute and State University*
Harvey Singer, *George Mason University*
Faye Teer, *James Madison University*

WASHINGTON

Sung Ahn, *Washington State University*
Stergio Fotopoulos, *Washington State University*
Lee McClain, *Western Washington University*

WASHINGTON D.C.

Rick Gibson, *American University*
Darius Singpurwalla, *George Washington University*
Ernest Zampelli, *Catholic University of America*

WISCONSIN

Patricia A. Mullins, *University of Wisconsin–Madison*
Cathy Poliak, *University of Wisconsin–Milwaukee*
Lee J. Van Scyoc, *University of Wisconsin–Oshkosh*

DEAR STUDENTS

Information overload is a recognized mark of our time, and the daily data tsunami isn't going to recede any time soon. There's no doubt you will be exposed to an unprecedented volume of data and information throughout your career. While many around you are crying "I'm overwhelmed!" or even "man overboard!" you can take a different, much better approach. Developing the skills needed to organize and interpret important information is a key to success, especially in business.

The business statistics course you are beginning will give you invaluable tools that enable you to use data to make good business decisions. For example, if you were a manager for AT&T or Verizon, could you conclude that the dropped-call rates for the two companies are significantly different based on a sample of customers from each company? Would that data alone be enough to affect how you manage your business? Or, would you want to break it down further and analyze it by some additional other factor, the locations of dropped calls perhaps? After you have completed Chapter 10, you'll be able to analyze data such as this and answer these questions with a high degree of reliability.

Throughout your business career you will find people at every level making decisions that have real consequences affecting business profitability and the jobs of real people. A skilled data analysis can give you an amazing window on aspects of a business that are simply not apparent without it.

We have written this textbook with you, the student, as our focus and the overall goal of helping you succeed both in this course and later in your career. The approach was developed over many years of teaching, and on the basis of that experience, We offer the following advice to help you achieve your own goals:

- Make it to class regularly. If you don't, you'll miss the details that help you master the subject. No textbook, no matter how well written, can take the place of your instructor and the classroom interaction. Seriously, go to class!
- Take advantage of the "Your Turn" problems placed throughout the chapters. Solving them will reinforce key concepts and let you know if you fully understand the material. The solutions to the problems at the end of each chapter give you immediate feedback, so I encourage you not to peek at the answers before you solve the problems or you won't really know how you are doing!
- Solve as many of the chapter problems as you can before exams. Business statistics is not a subject that rewards cramming or winging it. The saying "practice makes perfect" holds true in this field, and working through a variety of problems will build your confidence during the semester. We have provided solutions to all the even-numbered problems in Appendix B at the end of the book.

Make sure to use Excel as much as you can. Excel is important for students when entering the workplace as it is one of the most utilized business intelligence tools. We have provided detailed steps and figures on how to use Excel for statistical analysis.

We hope that you come to share our enthusiasm for the value of business statistics this semester and that what you learn in this course contributes to your future success in the business world.

INDEX OF APPLICATIONS

CHAPTER 1

An Introduction to Business Statistics

IN THIS CHAPTER, YOU WILL LEARN TO:

- How statistics is used in the business world.
- About the sources and data and the methods for collecting it.
- How to classify data by the level of measurement.
- To distinguish between time series and cross-sectional data.
- To distinguish between descriptive, inferential and predictive statistics.
- About the ethical implications of misusing statistics.

CHAPTER 1 MAP

Welcome to the world of statistics! Although some of you might be excited about learning statistics, others of you are probably less than thrilled about it. Perhaps you have to take "sadistics" because your major requires it. However, before you write off the value of this learning opportunity, let's discuss the role that statistics can play in your life. In today's world, everyone is a consumer of statistics. By this, I mean that you are continually surrounded by data and statements about those data in an effort to influence you to purchase something, vote for someone, or change your opinion about an issue. Consider the following examples:

- When CBS Sports announces that 103.4 million viewers tuned in to the 2018 Super Bowl (https://www.cbsnews.com/news/super-bowl-lii-tv-ratings/), do you understand the method used to determine this number? (How does CBS know that I, you, or anyone else watched the game?)
- When we hear on TV that President Donald Trump's approval rating is 40% and in small print see ±4%, do you understand the significance of this percentage?
- When you read an advertisement claiming that a new product is recommended by four out of five doctors, do you question the validity of the claim? (For instance, were the doctors paid for their endorsements?)
- When an online survey reports that Canon digital cameras are preferred to Nikon, does it concern you that the majority of the respondents could be Canon loyalists who repeatedly voted, skewing the results, or that the survey was conducted on a Canon user's forum?

Never before in the history of humankind have people had more data and information at their fingertips than you do at this moment. Statistics can have a powerful effect on our feelings, our opinions, and the decisions that we make in our personal and professional lives. As a result, it's very important that the statistics we report are both accurate and unbiased to ensure that they are properly utilized.

1.1 Business Statistics and Their Uses

Statistics is the mathematical science that deals with the collection, analysis, and presentation of data, which can then be used as a basis for inference and induction.

Business statistics are statistics applied to the business world to help improve decision making.

Statistics is the mathematical science that deals with the collection, analysis, and presentation of data—data that can then be used as a basis for inference and induction. **Business statistics** is the science of statistics applied to the business world in an effort to improve people's decision making in fields such as marketing, operations, finance, weather forecasting, and human resources, to name a few. Let's look at a few examples of how business statistics can help an organization's decision makers.

Marketing Research

Organizations rely heavily on business statistics when they conduct marketing research to determine what consumers want. For example, Kellogg's could perform a taste test to determine if consumers prefer the company's Cheez-It crackers to Nabisco's Cheese Nips. (Being a life-long Cheez-It addict, I know where my vote will go.) Kellogg's could also gather information about each consumer participating in the test in an effort to determine whether the people who prefer one brand to the other share similar characteristics. This would be useful information for Kellogg's future marketing efforts.

In the 1980s, Marriott conducted an extensive survey to see how potential customers felt about the company's current hotel offerings. Based on the results, Marriott designed a new hotel chain known as Courtyard by Marriott, which has been a huge success.

Advertising

Television networks set their advertising rates for commercials based on the sizes of viewing audiences. The networks receive the information from Nielsen Media Research, which collects data from approximately 41,000 U.S. households. The sample of households surveyed has been carefully selected so that the results can be used to infer the viewing habits of the entire country. Statistics are used to ensure the sample is properly chosen and to process the data into meaningful information for the networks. Based on this information, CBS was able to charge $4.5 million for a 30-second Super Bowl ad in 2018.

Operations

Statistics can also be used to help businesses operate better. Quality control, for example, is a vital concern for all successful organizations. The combination of business statistics and quality control is a marriage made in heaven. If properly implemented, statistics can help manufacturing and service organizations monitor their processes and determine when quality problems begin to occur. For instance, Kellogg's can use statistics to determine if my Cheez-Its are overbaked or too salty. (Based on the box I'm sampling from at this moment, I'd say the company is doing an excellent job with its statistical quality control.)

Finance

When Donnelly began writing this textbook, the U.S. economy was in the middle of a deep recession. Part of the economic downturn was due to the poor lending practices of banks, particularly in the mortgage industry. If used properly, business statistics is an excellent tool to help banks identify consumers who are good credit risks and those who are not, based on characteristics such as income, education, and home-ownership. For example, Fair Isaac Corporation (FICO) developed the credit scoring system (FICO score) currently used by the industry, which is based on a variety of statistical techniques.

Weather

The weather forecast helps many people in planning their decisions and activities. Statistics is necessary in forecasting the weather since the atmosphere is continuously changing and is not perfectly predictable. The use of predictive statistical

STATS IN PRACTICE: Careers in Statistics

It's a great time to be in the job market if you have an interest and aptitude for statistics. This is due to an increase in demand for the skills of individuals who have developed a level of statistical literacy. The biggest driver for this demand is a global society that is both data-rich and data-dependent as technology continues to advance at an explosive rate. In recognition of this development more than 150 professional organizations, including the American Statistical Association, designated 2013 as the International Year of Statistics.

There are abundant employment opportunities in today's business world that require expertise in statistics. The following are just a small sampling of the type of jobs available to people who have mastered this skill set.

- Marketing Analyst—Using statistics to analyze data related to people's purchasing patterns has helped businesses understand their consumers' buying behaviors. Companies that use data from social media platforms and create customized advertisements are examples of organizations that need this type of position to gain a competitive advantage.
- Financial analyst—Using a variety of statistical tools, financial analysts provide investment advice to businesses and individuals. Banks and investment firms are examples of organizations seeking this type of position.
- Actuary—In order to maintain profitability, insurance companies need actuaries to analyze risk factors for their customers in order to establish appropriate premiums for their service. Statistical techniques play a major role in this analysis. Hospitals, banks, and government agencies are other examples of organizations that rely on these skills.
- Sports statistician—Many professional baseball, basketball, and football teams have hired statisticians in an effort to gain a competitive advantage. Examples include:
  - Assistant General Manager Peter Brand, who was dramatized in the 2011 film *Moneyball*, used statistical analysis to help the Oakland Athletics assemble a competitive Major League Baseball team during the 2002 season. The Boston Red Sox won the World Series in 2004 (their first since 1918!) relying on some of the same statistical modeling used by the Oakland Athletics.
  - The Memphis Grizzlies, an NBA team, recently hired a vice president of basketball operations because of his statistical expertise.
- Political analyst—A great deal of attention has been paid to predicting political outcomes, such as elections, using statistical tools.
  - Nate Silver is an American statistician who has developed an impressive reputation for his accurate predictions in the political arena. I recommend visiting website, https://fivethirtyeight.blogs.nytimes.com/2013/, and reading his book *The Signal and the Noise: Why So Many Predictions Fail, But Some Don't*. Penguin, 2015 to gain some insight into the benefits of statistics in this field.

And of course I'd be remiss not mentioning one last type of employment in the field of statistics—education. I can personally testify that teaching statistics to students and writing books to help them learn has been a very rewarding experience.

models and technology forecasts the weather for a given location and time. Moreover, weather forecasts allows preparing for natural calamities, such as hurricanes and tornadoes to minimize the risk of disruptions and fatalities. For example, the weather forecast helped Floridians plan for Hurricane Irma in September 2017. Although many people had no power or water for days, the forecast allowed people to stock up on dry food and water many days before the hurricane hit.

1.2 Data

Data are values assigned to observations or measurements and are the building blocks of statistical analysis.

Data are the foundation of the field of statistics and can be defined as the values assigned to specific observations or measurements. If I'm collecting data on my wife's snoring behavior, I can do so in different ways. I can measure how many times Deb snores over a 10-minute period. I can measure the length of each snore in seconds. I could also measure how loud each snore is using a descriptive phrase like "That one sounded like a bear just waking up from hibernation" or "Wow! That one sounded like a sea lion calling for its young." (How a sound like that can come from a person who can fit into a pair of size 2 jeans is beyond me.)

In each instance, I would be recording data on the same event but in a different form. In the first instance, I would be measuring a frequency, or number of occurrences. In the second instance, I would be measuring duration, or length of time. And in the final instance, I would be measuring the event by describing its volume using words rather than numbers.

Information is data that are transformed into useful facts that can be used for a specific purpose, such as making a decision.

However, data all by themselves are not particularly useful. By definition, data are just the raw facts and figures that pertain to a measurement of interest. **Information**, on the other hand, is derived from the facts for the purpose of making decisions. One of the major reasons to use statistics is to transform data into information. For example, Table 1.1 shows the temperatures in New York City (NYC) during the first week of January 2018.

TABLE 1.1 | Temperatures in NYC in January 2018*

DATE	ACTUAL TEMP HIGH/LOW (°F)
Mon 1/1	19°/7°
Tue 1/2	26°/3°
Wed 1/3	30°/16°
Thu 1/4	29°/19°
Fri 1/5	19°/9°
Sat 1/6	13°/6°
Sun 1/7	18°/5°

*www.accuweather.com

Each individual temperature would be considered a data point. By themselves, the data points have limited value, other than to suggest that it is increasing or decreasing. To these data, let's add the fact that the lowest historical average temperature for the first week of January was 38°/27°. We can conclude now, with the help of statistics, that the temperatures were actually a record low for NYC. Adding another fact, that the freezing temperature was 32° Fahrenheit, we may be able to conclude now that people in NYC were at high risk of hypothermia (the body gets too cold and starts losing temperature at a higher rate than the body can generate). In order to address and analyze this, we will need to employ statistical analysis, which will be covered later in this text.

Data Sets and Databases

A **data sets** is a collection of data points in a single aspect.

A collection of data points is called a **data set**. Notice how Table 1.1 is a data set that shows a collection of temperature data points for one week. This data set has a single dimension for the seven days: actual temperature high/low.

A collection of data points that contains multiple rows (records) and columns (fields), and provides multiple dimensions for the data is called a **database**. Now consider the detailed weather forecast for the first week of January for NYC as shown in Table 1.2. The data is displayed in a multi-dimensional table that represents the forecast for multiple days (records) and multiple weather attributes (fields): day, actual temperature high/low, precipitation (chance of rain), snow, sunrise and sunset.

TABLE 1.2 | Weather Forecast Database for January 2018

DAY	ACTUAL TEMP HIGH/LOW (°F)*	PRECIP. (IN)*	SNOW (IN)*	SUNRISE**	SUNSET**
Mon 1/1	19°/7°	0	0	7:19 AM	4:39 PM
Tues 1/2	26°/3°	0	0	7:19 AM	4:40 PM
Wed 1/3	30°/16°	0	0	7:20 AM	4:41 PM
Thurs 1/4	29°/19°	0.76	0	7:20 AM	4:42 PM
Fri 1/5	19°/9°	0	0	7:20 AM	4:43 PM
Sat 1/6	13°/6°	0	0	7:19 AM	4:44 PM
Sun 1/7	18°/5°	0	0	7:19 AM	4:45 PM

Rows or Records

Columns or Fields

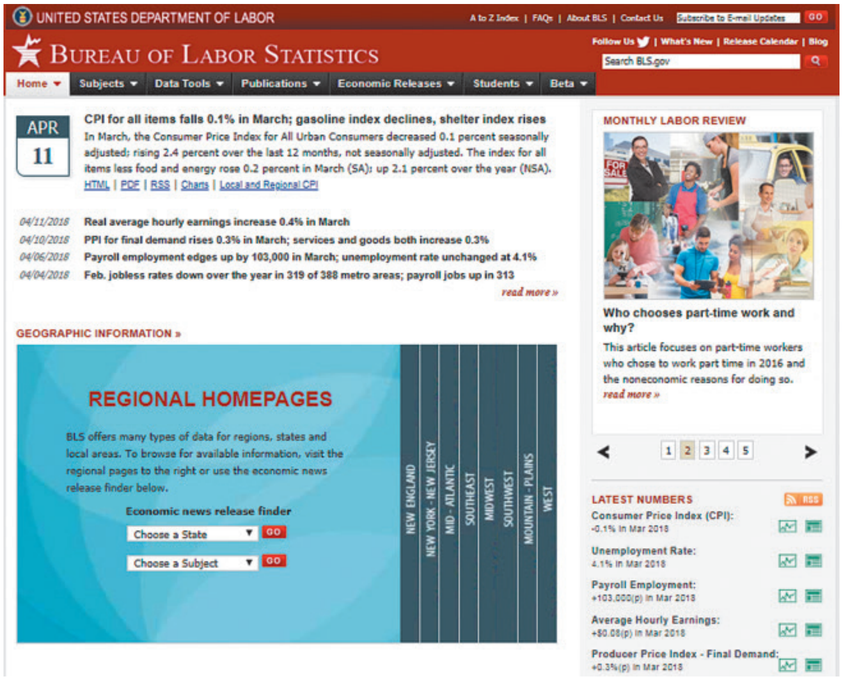
*www.accuweather.com
**www.weather.com

The Sources of Data

Secondary data are data that somebody else has collected and made available for others to use.

We classify the sources of data into two broad categories: primary data and secondary data. **Secondary data** are data somebody else has collected and made available for others to use. The U.S. government collects and publishes a variety of data that are readily available online. The U.S. Department of Labor collects mountains of data on topics such as consumer prices, inflation, unemployment, and productivity. The home page for the department's website is shown in Figure 1.1.

FIGURE 1.1 The Home Page of the Bureau of Labor Statistics



The Internet has become a rich source of data for statistics published by various industries. I once found a Japanese study on the effect of fluoride on toad embryos. Before this discovery, I had no idea toads even had teeth, much less a cavity problem!

Every 10 years the U.S. Department of Commerce conducts a nationwide census to gather a wide variety of data related to the country's population. The data are also used by Congress to make decisions about the funding for community services throughout the United States. The data are also used to adjust the number of representatives each state is allowed to elect to Congress. You can find census data on the Department of Commerce's website.

The U.S. Geological Survey (USGS) provides an impressive assortment of scientific information that is used to manage water, energy, biological, and mineral resources across the earth. For instance, did you know there are 250 species of squirrels in the world? If you don't believe me, you can look the information up at the USGS's website and become the local squirrel "expert" in your area.

The main drawback of using secondary data is that you have no control over how the data were collected. People tend to believe anything that's in print, even if it's not true. (You believe me, don't you?) Some of it is wrong, and, as you will learn later in the chapter, some of it is deliberately biased. The advantage to secondary data is that they are cheap (sometimes free) and that they are immediately available. For someone looking for data quickly, secondary data provide instant gratification (assuming the data are accurate, of course).

Primary data are data that you have collected for your own use.

Primary data, on the other hand, are data collected by the person or organization that eventually uses the data. This type of data can be expensive to acquire, but the main advantage of primary data is that the data are yours, and you have nobody else to blame but yourself if you make a mess of it. You can obtain primary data in many ways, such as by direct observation, via experiments, or through surveys.

Direct observation is a method of gathering data while the subjects of interest are in their natural environment, often unaware they are being watched.

Direct observation is a method of gathering data while the subjects of interest are in their natural environment, often unaware they are being watched. Observing wild animals stalking their prey in the forest or teenagers at the mall on Friday night are two examples. (Or are they the same example?) The advantage of this method is that the subjects will unlikely be influenced by the data collection process.

A **focus group** is a direct observational technique whereby individuals are often paid to discuss their attitudes toward products or services in a group setting controlled by a moderator.

A **focus group** is a direct observational technique whereby individuals are often paid to discuss their attitudes toward products or services in a group setting controlled by a moderator. For example, Fisher Price heavily relies on focus groups of both adults and children to obtain valuable feedback on new toy ideas. The participants are aware they are being observed.

An **experiment** is a method of gathering primary data by exposing subjects to certain treatments and recording the data of interest.

In an **experiment**, subjects are exposed to certain treatments and the data of interest are recorded. An experiment that tests the use of a new medical drug is an example. Two different groups would be established: The first group would receive the new drug; the second group would be the control group. People in the control group would be told they are getting the new drug but would in fact get a placebo with no medication. The reactions from each group would be measured and compared to determine whether the new drug is effective.

The benefit of experiments is that they allow statisticians to control factors that could influence the results, such as the gender, age, and education of a participant. One major concern about collecting data through experiments is that the response of the subjects might be influenced by the fact they are participating in a study.

A **survey** is a method of collecting primary data by directly asking people a series of questions. They can be administered by e-mail, via the Web, through snail mail, face to face, or over the telephone.

A **survey** involves directly asking people a series of questions. Surveys can be administered by e-mail, via the Web, through snail mail, face to face, or over the telephone. (It's the telephone survey I'm most fond of, especially when I get the call just as I'm sitting down to dinner, getting into the shower, or finally making some progress on the chapter I'm writing.) The questionnaire needs to be carefully designed to avoid any bias that could affect participants' responses or confuse them. Bias can occur when a question is stated in a way that encourages or leads a respondent to a particular answer. For example, "Wouldn't you agree that all drivers should wear a seat belt?" is a biased question. The influence the survey itself has on the responses of participants can also affect the quality of the data collected. Some participants will respond in a way they feel the survey would like them to. Figure 1.2 shows a portion of a survey I developed for users of the Claymont Community Center in Delaware. To encourage respondents to participate, an effective survey will state its purpose in the beginning, ask questions in a

Research has shown that a question posed in a positive tone will tend to evoke a more positive response. A question posed in a negative tone will tend to evoke a more negative response. A good strategy is to pre-test your questionnaire before releasing it to actual participants.

clear and concise manner, and place the more personal demographic questions last—when the respondents feel more comfortable with the process.

FIGURE 1.2
An Example of a Survey

Claymont Community Center
Health Clinic
Customer Service Survey

Dear Customer,

Our goal is to provide you with the best service possible. You can help us by taking a few minutes to answer the following questions about the Claymont Community Center (CCC). This valuable information will be used to help us to improve our services in the near future. The information gathered with this survey will be kept strictly confidential. Thank you for helping us.

(Please place this form in the Customer Survey Box in the Health Clinic. If you need more time, you may return the form on your next visit or ask for a stamped envelope to return it by mail.)

1. How long have you been using the Health Clinic? (Check one)
_____ Less than one month
_____ 1 to 6 months
_____ 6 months to a year
_____ 1 to 3 years

2. How satisfied are you with the quality of service being provided by the medical staff at the Health Clinic?
_____ Very Satisfied
_____ Satisfied
_____ Dissatisfied
_____ Very Dissatisfied
Comment____
____

3. Please provide the following personal information. Again, all information will remain confidential.

Zip code of your residence:_______________

Gender: _____ Male _____ Female Age: _____

Marital Status: _____ Married _____ Single
_____ Separated _____ Divorced _____ Widowed

Are you the head of your household?
_____ Yes _____ No

What is your annual gross income?
_____ Less than $10,000 _____ $30–40
_____ $10–20 _____ $40+
_____ $20–30

We relied on SurveyMonkey to develop this textbook. Faculty and students tested the book before publication and provided valuable feedback through this website. I'm a satisfied customer!

Online surveys are a convenient way to acquire data. Companies such as SurveyMonkey provide people with a low-cost way to design surveys, collect responses, and analyze the data. The SurveyMonkey website claims that 80% of the Fortune 100 companies are users of the service. However, there are challenges to Internet surveys, which will be discussed later in this chapter.

To test your understanding of data sources, I encourage you to spend a few minutes answering the questions in the following Your Turn section.

1 YOUR TURN

Identify if the data required for each example are primary or secondary. For primary data, determine the best way in which the data should be collected. In other words, should the data be collected via observation, experiment, or survey?

1. Apple would like to measure the satisfaction levels of customers who purchased its new iPad product.
2. Pepsi would like to determine if consumers prefer the taste of Diet Pepsi to Diet Coke.
3. Cleveland State University needs to determine the current inflation rate to determine the annual salary increases for its staff for the upcoming year.
4. McDonald's would like to determine the average wait time for customers who use its drive-through windows during the lunch hour.

Answers can be found on **page 21.**

Quantitative data use numerical values while **qualitative data** rely on descriptive terms to describe something of interest.

Quantitative continuous data are values that can take on any real numbers, including numbers that contain decimal points.

Quantitative discrete data are values based on observations that can be counted and are typically represented by whole numbers.

The Two Main Types of Data

Another way to classify data is by whether they are quantitative or qualitative:

- **Quantitative data** use numerical values to describe something of interest either by measuring it (such as its weight, height, or distance), we call this **continuous data,** or by counting it (such as the number of customers or repeat customers a business receives)—we call this **discrete data.**

- **Qualitative data** use descriptive terms to measure or classify something of interest. One example of qualitative data is the name of a respondent in a survey and his or her level of education. Mathematical operations such as addition, subtraction, multiplication, and division cannot be performed on this type of data.

Classifying Data by Their Level of Measurement

Another important way to classify data is by the *way* the data are measured. This distinction is critical because it affects which statistical techniques we can use in our analysis of the data. The four levels of measurement are nominal, ordinal, interval, and ratio.

Nominal data are data that are described as a category or a label. Examples are gender (male or female), marital status (married, single, divorced, widowed), or yes/no responses.

A nominal level of measurement deals strictly with qualitative data assigned to predetermined categories. One example is the gender of a survey respondent, with the categories being male and female. This type of data is referred to as **nominal data,** or categorical data. It does not allow us to perform any mathematical operations on it, such as adding or multiplying. We can only give the data names and categorize them. (The word *nominal* actually means "pertaining to names.") We also cannot rank-order the data in any way from highest to lowest. An example is the state in which the survey respondent resides, such as Delaware or New Jersey (although I would try to rank my home state of Delaware on top.)

Other examples of nominal data are zip codes and telephone numbers, which can't be added, subtracted, or placed in a meaningful order of greater than or less than. Even though the data consist of numbers, they are handled just like qualitative data. Nominal data are considered the lowest level of data, and, as a result, the statistical techniques used to analyze them are the most restrictive.

Ordinal data have all the properties of nominal data, with the added feature that we can rank-order the values from highest to lowest.

An ordinal level of measurement can be conducted on data that are on the next level up on the food chain. **Ordinal data** has all the properties of nominal data, with the added feature that we can rank-order the values from highest to lowest. The following example explains ordinal measurement: Recently, I felt challenged by two neighbors who claimed their lawnmowers were faster than mine. Naturally, this had to be settled by a lawnmower race down our street, Gaebel Lane. Sadly, I present Figure 1.3, which shows that I lost the race. My neighbor Tom came in first, my neighbor Scott came in second, and I came in third.

FIGURE 1.3
An Example of Ordinal Measurement: Tom, Me, and Scott (from left to right) on Our Mowers

We still can't perform mathematical operations on this type of data, but we can say that Tom's lawnmower was faster than mine and Scott's. However, we cannot say how *much* faster because we didn't record the times of the lawnmowers. We just noted who came in first, second, and third. Ordinal data that have been collected do not allow us to make measurements between the categories or to say, for instance, that Scott's lawnmower is twice as fast as Bob's. For that, we need a different type of data. (In case you were wondering, I have been unsuccessful at convincing Deb that I need more horsepower to restore our family honor at the Second Annual Gaebel Lane Lawnmower Race.)

Education level is another example of ordinal data. A master's degree is ranked higher than a bachelor's degree, which in turn is ranked higher than a high school diploma. However, we are unable to measure the difference between these degrees in a meaningful, mathematical way. For instance, it would not be accurate to claim the difference between a master's and a bachelor's degree is more than the difference between a bachelor's and a high school degree. A property of ordinal data is that the differences between categories are not meaningful and, therefore, cannot be measured.

Ordinal data can also be numerical. One example of numerical ordinal data is when we rate movies with one, two, three, or four stars. Although we can order the movies by their ratings, we can't, for example, claim that a four-star movie is four times as good as a one-star movie.

Interval data, which are strictly quantitative, allow us to measure the differences between the categories with actual numbers in a meaningful way.

The interval measurement level is yet a higher level of measurement. It measures **interval data,** which are strictly quantitative. Temperature measurements in degrees Fahrenheit are an example of interval data. With this level, we can measure the differences between the categories with actual numbers in a meaningful way. For instance, 70°F is 5° warmer than 65°F. However, multiplication and division can't be performed on this level of data. Why not? Simply because we cannot argue that 100°F is twice as warm as 50°F. The logic of this claim becomes more obvious when we convert the temperatures to the Celsius scale. The same two temperatures convert to 38°C and 10°C, respectively, so the twice-as-warm argument does not hold true. To help explain this, try baking a cake at twice the recommended temperature in half the recommended time. Yuck!

Another characteristic of interval data is that they do not have a true zero point. The term *true zero point* means that a zero data value indicates the absence of the object being measured. For instance, 0°F and 0°C do not represent the absence of temperature, even though it may feel like it.

Your grade point average (GPA) is another example of interval data. We can measure the difference between a 4.0 and a 2.0 GPA by simply subtracting the two values. However, it would not be an accurate statement to claim that a 4.0 student is twice as smart as a 2.0 student. Also, GPA has no true zero point because a 0.0 GPA does not indicate the absence of a grade point average.

Ratio data have all the features of interval data, with the added benefit of having a true zero point.

The most versatile of data types is the ratio level of measurement. **Ratio data** are as good as it gets as far as data are concerned. Examples of this type of data are ages, weights, prices, and salaries. Ratio data have all the features of interval data, with the added benefit of a true zero point. For instance, a zero salary indicates the absence of any salary. With a true zero point, we can say that a person who is six feet in height is twice as tall as a three-foot person or that a 20-year-old person is half the age of a 40-year-old.

For example, dollars are considered to be ratio data because \$20 is twice as much as \$10.

The distinction between interval and ratio data is a fine line. To help identify the proper scale, use the "twice as much" rule. If the phrase "twice as much" accurately describes the relationship between two values that differ by a multiple of two, then the data can be considered to be ratio level.

Figure 1.4 shows the relationship between the levels of data measurement and the two main types of data, and Table 1.3 summarizes the properties of the four levels of data measurement. As we explore different statistical techniques later in this book, we will revisit these different measurement scales. You will discover that specific techniques require certain types of data.

FIGURE 1.4 The Two Main Types of Data and Their Corresponding Levels